

Abstract Submitted  
for the MAR14 Meeting of  
The American Physical Society

**Geometrical structure of Neural Networks: Geodesics, Jeffrey's Prior and Hyper-ribbons**<sup>1</sup> LORIEN HAYDEN, ALEX ALEMI, JAMES SETHNA, Cornell University — Neural networks are learning algorithms which are employed in a host of Machine Learning problems including speech recognition, object classification and data mining. In practice, neural networks learn a low dimensional representation of high dimensional data and define a model manifold which is an embedding of this low dimensional structure in the higher dimensional space. In this work, we explore the geometrical structure of a neural network model manifold. A Stacked Denoising Autoencoder and a Deep Belief Network are trained on handwritten digits from the MNIST database. Construction of geodesics along the surface and of slices taken from the high dimensional manifolds reveal a hierarchy of widths corresponding to a hyper-ribbon structure. This property indicates that neural networks fall into the class of sloppy models, in which certain parameter combinations dominate the behavior. Employing this information could prove valuable in designing both neural network architectures and training algorithms.

<sup>1</sup>This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No . DGE-1144153.

Lorien Hayden  
Cornell University

Date submitted: 14 Nov 2013

Electronic form version 1.4